

WO 2004/055189

1/8

SEQUENCE LISTING

<110> Akzo Nobel N.V.

<120> New expression system from Rhodococcus

<130>

<140>

<141>

<160> 13

<170> PatentIn·Ver. 2.1

<210> 1

<211> 1543

<212> DNA

<213> Rhodococcus erythrop olis

<220>

<221> CDS

<222> (1)...(1533)

<400> 1

atg cag gac tgg acc agc gag tgc gac gtg ttg gta gtc ggc tcc ggc 48  
Met Gln Asp Trp Thr Ser Glu Cys Asp Val Leu Val Val Gly Ser Gly  
1 5 10 15

ggc gga gcg ctg acc ggc gca tat acc gcc gct gct cag gga ttg acg 96  
Gly Gly Ala Leu Thr Gly Ala Tyr Thr Ala Ala Gln Gly Leu Thr  
20 25 30

acg atc gtc ctc gag aaa acc gat cgt ttc ggc ggg acc tcc gc c tac 144  
Thr Ile Val Leu Glu Lys Thr Asp Arg Phe Gly Gly Thr Ser Ala Tyr  
35 40 45

tcg ggc gcc tcg atc tgg ctc cca ggt acc cag gtg cag gaa cgc gcc 192  
Ser Gly Ala Ser Ile Trp Leu Pro Gly Thr Gln Val Gln Glu Arg Ala  
50 55 60

gga ctt ccc gac tcg acc gag aat gcc cgc acc tat ctg cgc gcg ttg 240  
Gly Leu Pro Asp Ser Thr Glu Asn Ala Arg Thr Tyr Leu Arg Ala Leu  
65 70 75 80

ctc ggt gac gcc gag tcc gag cgc cag gac gcc tac gtc gag acc gct 288  
Leu Gly Asp Ala Glu Ser Glu Arg Gln Asp Ala Tyr Val Glu Thr Ala  
85 90 95

ccc gct gtc gtc gct cta ctc gag cag aac ccg aac atc gaa ttc gag 336  
Pro Ala Val Val Ala Leu Leu Glu Gln Asn Pro Asn Ile Glu Phe Glu  
100 105 110

ttc cgt gcg ttc ccc gac tac tac aaa gcc gaa ggc cgg atg gac acg 384  
Phe Arg Ala Phe Pro Asp Tyr Tyr Lys Ala Glu Gly Arg Met Asp Thr  
115 120 125

gga cgc tcc atc aac cct ctc gat ctc gat ccc gcc gac atc ggt gac 432  
Gly Arg Ser Ile Asn Pro Leu Asp Leu Asp Pro Ala Asp I le Gly Asp  
130 135 140

ctc gcc ggc aag gtg cgt ccg gaa ctg gac caa gac cgc acc ggt cag 480  
Leu Ala Gly Lys Val Arg Pro Glu Leu Asp Gln Asp Arg Thr Gly Gln  
145 150 155 160

gat cat gct ccc ggc ccg atg atc ggt ggg cgc gca ctg atc ggc cgt 528  
Asp His Ala Pro Gly Pro Met Ile Gly Arg Ala Leu Ile Gly Arg  
165 170 175

ctg ctc gcc gca gtt cag agc acc ggt aag gca gaa ctt cgc acc gaa Leu Leu Ala Ala Val Gln Ser Thr Gly Lys Ala Glu Leu Arg Thr Glu 180	185	190	576
tcc gtc ctc acc tcc ctg atc gtg gaa gac ggc cgt gtt gtc ggc gcc Ser Val Leu Thr Ser Leu Ile Val Glu Asp Gly Arg Val Val Gly Ala 195	200	205	624
gag gtc gaa tcc ggc ggc gaa acc cag cga atc aag gcg aac cgc ggt Glu Val Glu Ser Gly Gly Glu Thr Gln Arg Ile Lys Ala Asn Arg Gly 210	215	220	672
gtc ctg atg gca gca ggc ggc atc gaa ggc aac gcc gag atg cgt gag Val Leu Met Ala Ala Gly Gly Ile Glu Gly Asn Ala Glu Met Arg Glu 225	230	235	720
cag gca ggc acc ccc ggc aag gcg atc tgg agt atg ggt ccc ttc ggc Gln Ala Gly Thr Pro Gly Lys Ala Ile Trp Ser Met Gly Pro Phe Gly 245	250	255	768
gcc aac acc ggc gac gcg atc tct gcc ggt att gc t gtc ggc ggc gca Ala Asn Thr Gly Asp Ala Ile Ser Ala Gly Ile Ala Val Gly Gly Ala 260	265	270	816
aca gcc ttg ctc gat cag gcg tgg ttc tgc ccc ggc gtc gag cag ccc Thr Ala Leu Leu Asp Gln Ala Trp Phe Cys Pro Gly Val Glu Gln Pro 275	280	285	864
gac ggc agc gcc gcc ttc atg gtc ggc gtt cgc ggt ggg ctc gtc gtc Asp Gly Ser Ala Ala Phe Met Val Gly Val Arg Gly Gly Leu Val Val 290	295	300	912
gac agc gcc ggt gag cgc tac ctc aac gag tcg ctt ccg tac gac cag Asp Ser Ala Gly Glu Arg Tyr Leu Asn Glu Ser Leu Pro Tyr Asp Gln 305	310	315	960
ttc gga cga gcc atg gat gct cac gac gac aac ggt tct gcc gtg ccg Phe Gly Arg Ala Met Asp Ala His Asp Asp Asn Gly Ser Ala Val Pro 325	330	335	1008
tcg ttc atg atc ttc gac tcg cgc gag ggt ggc gga ctg ccc gcc atc Ser Phe Met Ile Phe Asp Ser Arg Glu Gly Gly Leu Pro Ala Ile 340	345	350	1056
tgc atc ccg aac acg gcg ccc gcc aag cac ctc gaa gcc gga acg tgg Cys Ile Pro Asn Thr Ala Pro Ala Lys His L eu Glu Ala Gly Thr Trp 355	360	365	1104
gtc ggt gcc gac act ctc gaa gaa ctc gct gcc aag acc gga cta ccg Val Gly Ala Asp Thr Leu Glu Glu Leu Ala Ala Lys Thr Gly Leu Pro 370	375	380	1152
gcc gac gca ttg cgc agc act gtc gaa aag ttc aac gat gcc gca aaa Ala Asp Ala Leu Arg Ser Thr Val Glu Lys Phe Asn Asp Ala Ala Lys 385	390	395	1200
ctg ggc gtc gac gaa gag ttc cat cgc ggc gaa gac cog tac gac gcg Leu Gly Val Asp Glu Glu Phe His Arg Gly Glu Asp Pro Tyr Asp Ala 405	410	415	1248
ttc ttc tgc cca ccc aac ggc ggt gcg aat gcg gca ctg acg gcc atc Phe Phe Cys Pro Pro Asn Gly Gly Ala Asn Ala Ala Leu Thr Ala Ile 420	425	430	1296
gag aac gga ccg ttc tac gcg gcc cgc atc gtc ctc agt gac ctc ggc Glu Asn Gly Pro Phe Tyr Ala Ala Arg Ile Val Leu Ser Asp Leu Gly 445	450	455	1344

435	440	445	
acc aag ggc gga ttg gtc acc gac gtc aac ggc cga gtc ctg cgt gct Thr Lys Gly Gly Leu Val Thr Asp Val Asn Gly Arg Val Leu Arg Ala 450 455 460			1392
gac ggc agc gcc atc gac ggc ctg tac gcc gcc ggc aac acg acg gcg Asp Gly Ser Ala Ile Asp Gly Leu Tyr Ala Ala Gly Asn Thr Ser Ala 465 470 475 480			1440
tca ctg agc ggc cgc ttc tac ccc gg c ccc gga gtt cca ctc ggc acg Ser Leu Ser Gly Arg Phe Tyr Pro Gly Pro Gly Val Pro Leu Gly Thr 485 490 495			1488
gct atg gtc ttc tcg tac cga gca gct cag gac atg gcg aag taa Ala Met Val Phe Ser Tyr Arg Ala Ala Gln Asp Met Ala Lys 500 505 510			1533
cgcagttcaa			1543
 <b>&lt;210&gt; 2</b> <b>&lt;211&gt; 511</b> <b>&lt;212&gt; PRT</b> <b>&lt;213&gt; Rhodococcus erythropolis</b>			
 <b>&lt;400&gt; 2</b> Met Gln Asp Trp Thr Ser Glu Cys Asp Val Leu Val Val Gly Ser Gly 1 5 10 15			
Gly Gly Ala Leu Thr Gly Ala Tyr Thr Ala Ala Ala Gln Gly Leu Thr 20 25 30			
Thr Ile Val Leu Glu Lys Thr Asp Arg Phe Gly Gly Thr Ser Ala Tyr 35 40 45			
Ser Gly Ala Ser Ile Trp Leu Pro Gly Thr Gln Val Gln Glu Arg Ala 50 55 60			
Gly Leu Pro Asp Ser Thr Glu Asn Ala Arg Thr Tyr Leu Arg Ala Leu 65 70 75 80			
Leu Gly Asp Ala Glu Ser Glu Arg Gln Asp Ala Tyr Val Glu Thr Ala 85 90 95			
Pro Ala Val Val Ala Leu Ieu Glu Gln Asn Pro Asn Ile Glu Phe Glu 100 105 110			
Phe Arg Ala Phe Pro Asp Tyr Tyr Lys Ala Glu Gly Arg Met Asp Thr 115 120 125			
Gly Arg Ser Ile Asn Pro Leu Asp Leu Asp Pro Ala Asp Ile Gly Asp 130 135 140			
Leu Ala Gly Lys Val Arg Pro Glu Leu Asp Gln Asp Arg Thr Gly Gln 145 150 155 160			
Asp His Ala Pro Gly Pro Met Ile Gly Gly Arg Ala Leu Ile Gly Arg 165 170 175			
Leu Leu Ala Ala Val Gln Ser Thr Gly Lys Ala Glu Leu Arg Thr Glu 180 185 190			
Ser Val Leu Thr Ser Leu Ile Val Glu Asp Gly Arg Val Val Gly Ala 195 200 205			
Glu Val Glu Ser Gly Glu Thr Gln Arg Ile Lys Ala Asn Arg Gly			

210	215	220
Val Leu Met Ala Ala Gly	Gly Ile Glu Gly	Asn Ala Glu Met Arg
225	230	235 240
Gln Ala Gly Thr Pro	Gly Lys Ala Ile Trp Ser Met	Gly Pro Phe Gly
245	250	255
Ala Asn Thr Gly Asp Ala Ile Ser Ala	Gly Ile Ala Val	Gly Ala
260	265	270
Thr Ala Leu Leu Asp Gln Ala Trp Phe Cys Pro	Gly Va 1 Glu Gln Pro	
275	280	285
Asp Gly Ser Ala Ala Phe Met Val	Gly Val Arg Gly	Gly Leu Val Val
290	295	300
Asp Ser Ala Gly Glu Arg Tyr Leu Asn Glu	Ser Leu Pro	Tyr Asp Gln
305	310	315 320
Phe Gly Arg Ala Met Asp Ala His Asp Asp Asn	Gly Ser Ala Val	Pro
325	330	335
Ser Phe Met Ile Phe Asp Ser Ar g Glu	Gly Gly Leu Pro	Ala Ile
340	345	350
Cys Ile Pro Asn Thr Ala Pro Ala Lys His Leu	Glu Ala Gly Thr Trp	
355	360	365
Val Gly Ala Asp Thr Leu Glu Glu Leu Ala Ala	Lys Thr Gly Leu Pro	
370	375	380
Ala Asp Ala Leu Arg Ser Thr Val Glu Lys	Phe Asn Asp Ala Ala	Lys
385	390	395 400
Leu Gly Val Asp Glu Glu Phe His Arg Gly	Glu Asp Pro	Tyr Asp Ala
405	410	415
Phe Phe Cys Pro Pro Asn Gly Gly Ala Asn Ala Ala	Leu Thr Ala Ile	
420	425	430
Glu Asn Gly Pro Phe Tyr Ala Ala Arg Ile Val	Leu Ser Asp Leu	Gly
435	440	445
Thr Lys Gly Gly Leu Val Thr Asp Val Asn Gly	Arg Val Leu Arg	Ala
450	455	460
Asp Gly Ser Ala Ile Asp Gly Leu Tyr Ala Ala	Gly Asn Thr Ser	Ala
465	470	475 480
Ser Leu Ser Gly Arg Phe Tyr Pro Gly Pro	Gly Val Leu	Gly Thr
485	490	495
Ala Met Val Phe Ser Tyr Arg Ala Ala Gln Asp	Met Ala Lys	
500	505	510

&lt;210&gt; 3

&lt;211&gt; 158

&lt;212&gt; DNA

&lt;213&gt; Rhodococcus erythropolis

&lt;400&gt; 3 atctatcgatt atgtgtcccg gcccgcgaacg a cccgcgtaa ttctctcacc tggaccaccc 60

atctcggcat attgccccgt cagtgggacc tggcatggcc ttccagtgcc gtgcggatt 120

ccgtggacac cccaccctct tggagtaagg acgcaatg	158
<210> 4	
<211> 19	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Description of Artificial Sequence:primer	
<400> 4	
ggcgacgttgc cgagaatt	19
<210> 5	
<211> 624	
<212> DNA	
<213> Rhodococcus erythropolis	
<220>	
<221> CDS	
<222> (1)..(624)	
<400> 5	
atg ggg gcg acg ttg ccg aga att gcc gag gtc agg gac gct gct gag	48
Met Gly Ala Thr Leu Pro Arg Ile Ala Glu Val Arg Asp Ala Ala Glu	
1 5 10 15	
ccc agt tcg gac gag cag cgg gcg cgc cat gtg cgg atg ctg gaa gcg	96
Pro Ser Ser Asp Glu Gln Arg Ala Arg His Val Arg Met Ieu Glu Ala	
20 25 30	
gcc gcc gaa ttg ggg acc gag aaa gaa ctc tca ccg gtt cag atg cac	144
Ala Ala Glu Leu Gly Thr Glu Lys Glu Leu Ser Arg Val Gln Met His	
35 40 45	
gaa gtt gcc aag cgg gca ggc gtg gcc atc ggc act ctc tac cgc tat	192
Glu Val Ala Lys Arg Ala Gly Val Ala Ile Gly Thr Leu Tyr Arg Tyr	
50 55 60	
tgc cct tcg aag acg cac ctc ttc gtc gct gtg atg gtc gag cag atc	240
Phe Pro Ser Lys Thr His Leu Phe Val Ala Val Met Val Glu Gln Ile	
65 70 75 80	
gat cag atc ggc gac agt ttc gcc aag cat c ag gtg cag tcg gcc aat	288
Asp Gln Ile Gly Asp Ser Phe Ala Lys His Gln Val Gln Ser Ala Asn	
85 90 95	
ccg cag gac gcc gtg tac gag gtc ctg gtg cgc gcg act cgc ggg tta	336
Pro Gln Asp Ala Val Tyr Glu Val Leu Val Arg Ala Thr Arg Gly Leu	
100 105 110	
ctg cgt cgg ccg gcc ctt tcg act gcg atg ctg cag tcg tcc agt acc	384
Leu Arg Arg Pro Ala Leu Ser Thr Ala Met Leu Gln Ser Ser Thr	
115 120 125	
gcc aac gtc gcg acg gtg ccg gac gtg ggc aag atc gat cgc ggc ttc	432
Ala Asn Val Ala Thr Val Pro Asp Val Gly Lys Ile Asp Arg Gly Phe	
130 135 140	
cgg cag atc atc ctc gat gcg gcc ggg atc gag aac ccg acc gag gaa	480
Arg Gln Ile Ile Leu Asp Ala Ala Gly Ile Glu Asn Pro Thr Glu Glu	
145 150 155 160	
gac aac acc ggg ttg cgt ctg atg cag ctg tgg ttc ggg gtc atc	528
Asp Asn Thr Gly Leu Arg Leu Leu Met Gln Leu Trp Phe Gly Val Ile	

165

170

175

caa tcg tgc ctc aac ggt cga att tcc atc ccg gat gcg gag tac gac 576  
 Gln Ser Cys Leu Asn Gly Arg Ile Ser Ile Pro Asp Ala Glu Tyr Asp  
 180 185 190

atc cgc aag ggg tgc gac ctg ctt ctg gtg aat ctc tca cga cac tga 624  
 Ile Arg Lys Gly Cys Asp Leu Leu Val Asn Leu Ser Arg His  
 195 200 205

<210> 6  
 <211> 208  
 <212> PRT  
 <213> Rhodococcus erythropolis

<400> 6  
 Met Gly Ala Thr Leu Pro Arg Ile Ala Glu Val Arg Asp Ala Ala Glu  
 1 5 10 15

Pro Ser Ser Asp Glu Gln Arg Ala Arg His Val Arg Met Leu Glu Ala  
 20 25 30

Ala Ala Glu Leu Gly Thr Glu Lys Glu Leu Ser Arg Val Gln Met His  
 35 40 45

Glu Val Ala Lys Arg Ala Gly Val Ala Ile Gly Thr Leu Tyr Arg Tyr  
 50 55 60

Phe Pro Ser Lys Thr His Leu Phe Val Ala Val Met Val Glu Gln Ile  
 65 70 75 8 0

Asp Gln Ile Gly Asp Ser Phe Ala Lys His Gln Val Gln Ser Ala Asn  
 85 90 95

Pro Gln Asp Ala Val Tyr Glu Val Leu Val Arg Ala Thr Arg Gly Leu  
 100 105 110

Leu Arg Arg Pro Ala Leu Ser Thr Ala Met Leu Gln Ser Ser Ser Thr  
 115 120 125

Ala Asn Val Ala Thr Val Pro Asp Val Gly Lys Ile Asp Arg Gly Phe  
 130 135 140

Arg Gln Ile Ile Leu Asp Ala Ala Gly Ile Glu Asn Pro Thr Glu Glu  
 145 150 155 160

Asp Asn Thr Gly Leu Arg Leu Leu Met Gln Leu Trp Phe Gly Val Ile  
 165 170 175

Gln Ser Cys Leu Asn Gly Arg Ile Ser Ile Pro Asp Ala Glu Tyr Asp  
 180 185 190

Ile Arg Lys Gly Cys Asp Leu Leu Val Asn Leu Ser Arg His  
 195 200 205

<210> 7  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:primer

<400> 7  
 tcagtgtcgt gagagattca

<210> 8  
<211> 35  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:primer

<400> 8  
ataaaagctta tcgattatgt gtcccgccg cgaac 35

<210> 9  
<211> 35  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:primer

<400> 9  
ataggtacca tatgtgcgtc cttaactccaa gaggg 35

<210> 10  
<211> 27  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:primer

<400> 10  
gcgcatatgg ctaagaatca g gcaccc 27

<210> 11  
<211> 30  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:primer

<400> 11  
gcgggatccc tacttctctg ctgcgtgatg 30

<210> 12  
<211> 30  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:primer

<400> 12  
ggccatatgt tgaccacaga cgtgacgacc 30

<210> 13  
<211> 30  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:primer

<400> 13  
gccactagtt cactgcgctg ctctgcacg 30